

#112
03-d-d

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/508,821A
 DATE: 02/23/2001
 TIME: 16:03:53

Input Set : A:\Pto.amc
 Output Set: N:\CRF3\02232001\I508821A.raw

```

2 <110> APPLICANT: ROULEAU, Guy A.
3 JOOBER, Ridha
4 BENKELFAT, Chawki
6 <120> TITLE OF INVENTION: POLYMORPHIC CAG REPEAT-CONTAINING GENE AND USES THEREOF
8 <130> FILE REFERENCE: 12948-10
10 <140> CURRENT APPLICATION NUMBER: 09/508,821A
11 <141> CURRENT FILING DATE: 2000-05-26
13 <150> PRIOR APPLICATION NUMBER: PCT/CA98/00884
14 <151> PRIOR FILING DATE: 1998-09-18
16 <150> PRIOR APPLICATION NUMBER: 2,216,057 CA
17 <151> PRIOR FILING DATE: 1997-09-19
19 <160> NUMBER OF SEQ ID NOS: 6
21 <170> SOFTWARE: PatentIn Ver. 2.1
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 328
25 <212> TYPE: DNA
26 <213> ORGANISM: Mus musculus
28 <400> SEQUENCE: 1
29 agggcagcca ctttcccag cattcccagt ccttccctac ctccctccact tatgcccaca 60
30 cagtgcaggg tgggtggcag ggggcccact cctacaagag ctgcacagca ccatctgccc 120
31 agcctcatga taggcgatg agtgccaatg cgaacctggc tccagggcaa cgggtccaga 180
32 atcttccagc ttaccagcct ggcgccttg gctacgagca gcagcagcaa gcacttccag 240
33 gccgtacca caccaggaa acactccact accagaaact cgcgaagtac caacactatg 300
34 gacagcaagg ccagggtac tgtccacc 328
37 <210> SEQ ID NO: 2
38 <211> LENGTH: 330
39 <212> TYPE: DNA
40 <213> ORGANISM: Homo sapiens
42 <220> FEATURE:
43 <221> NAME/KEY: modified_base
44 <222> LOCATION: (59)
45 <223> OTHER INFORMATION: N = A or G or C or T/U
47 <400> SEQUENCE: 2
W--> 48 tccttcccca cctcctccac ctactcctcc tctgtcccag ggtggtgggc agggggccna 60
49 ctctataag agttgcacag caccgactgc ccagcccact gacaggccgc tgactgccag 120
50 ctccagcctg gccccggggc agcgggtcca gaatcttcat gctaccagc cgggcgcct 180
51 cagctatgac cagcagcagc agcagcagca gcagcagcag cagcagcagc aagcccttca 240
52 gagccggcac catgcccagg aaaccctcca ttaccaaaac ctgcgaagt atcagcacta 300
53 cgggcagcaa ggcagggct actgccagcc 330
56 <210> SEQ ID NO: 3
57 <211> LENGTH: 23
58 <212> TYPE: DNA
59 <213> ORGANISM: Artificial Sequence
61 <220> FEATURE:
62 <223> OTHER INFORMATION: Description of Artificial Sequence:
63 oligonucleotide
65 <400> SEQUENCE: 3

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/508,821A

DATE: 02/23/2001
 TIME: 16:03:53

Input Set : A:\Pto.amc
 Output Set : N:\CRF3\02232001\I508821A.raw

```

66 ggggcagcgg gtccagaalc ttc
69 <210> SEQ ID NO: 4
70 <211> LENGTH: 23
71 <212> TYPE: DNA
72 <213> ORGANISM: Artificial Sequence
74 <220> FEATURE:
75 <223> OTHER INFORMATION: Description of Artificial Sequence:
76     oligonucleotide
78 <400> SEQUENCE: 4
79 agcactacgg gcagcaaggg cag
82 <210> SEQ ID NO: 5
83 <211> LENGTH: 6022
84 <212> TYPE: DNA
85 <213> ORGANISM: Homo sapiens
87 <400> SEQUENCE: 5
88 ggatccagca ggcacaaagg gatgaggag cggaattgc tctgctaaat gcttttgagc 60
89 tglcaggaaq ggtgggagt gatgggtggg ggacattggg gaggagctgg caatgggagg 120
90 gggggggggc ggtagctccc cagtgcctgg gcgctgggca gccggttttg cctcccgcat 180
91 cagtggccgt ccttgccaag actcagctgc aggcgatgtg ggagcgggaa ttacagagca 240
92 caectccctg acacagaagt tglcaatatg cgcacaqctg ggggggaggg ccaggcgaag 300
93 ggggggactat taagagctgc gggggggagg aggcagggtg gggaggtggg tgggaggttg 360
94 cttctgagg caaaagggaag tggcccgctc gaatcgctca tctctgccc cctccctgcc 420
95 catctctccc tctctctccc cctccctccc tctctctctt tttcttlica cagataacca 480
96 gcccgagtca tgcagttctt tggagaaaag tglggtttcc atggcaaaac acagaactac 540
97 cagcagacct cgcaggaaac atcacgccta gagaattaca ggcagccgag ccaggccggg 600
98 ctaagctgcg accggcagcg gctgctcgcc aaggactatt ataaccggca gcttaccggg 660
99 agctatgagg gtggcgctgg caagccctct ggcactgcag ccgcggtggc cgcgcacaag 720
100 taccacccag gcagcaaggc cctgcccaca cagcaaggcc tgcaggggag gccggcttcc 780
101 cctggctacg gcgcccagga cagcagcccc taaccaggcc gctatgctgg tgaggagagg 840
102 ctccaggctt gggggggccc acagccacca ccccacagc cgcagccact acctgcaggg 900
103 gtggccaaqt atgatgagaa ctgatgaaa aagacagcag tgcacccagc caggcaglat 960
104 gcagagcagq gcqcccaggt gcccttttcg actcactccc tgcacgtcca gcagccaccg 1020
105 ccgccccagc agccctctgg ataccccaag ctocaaaggg agaagctgca gaacgacatt 1080
106 gccctcccctc tgcctctccc ccagggtacc cactttcttc agcatttccc gtccttccc 1140
107 acctctctca cctactctc ctctgtccag ggtgggtggc agggggccca ctctataag 1200
108 agttgcacag caccgactgc ccagccccat gacaggccgc tgactgccag ctccagcctg 1260
109 gccccggggc agcgggtcca gaattctcat gcctaccagt cgggcccctt cagctatgac 1320
110 cagcagcagc agcagcagca gcagcagcag cagcagcagc aagcccttca gagccggcac 1380
111 catgccagg aaaccttcca ttaccaaaac ctgcgaagt atcagcacta cgggcagcaa 1440
112 ggcagggct actgccagcc ggacgcagcc gtcgggaccc cagagcagta ctaccagacc 1500
113 ttccagccca gctccagcca ctacccggc cgtctcgtgg gccgtcacc ttctacagt 1560
114 tccacaccgt cgcctgcat gccaaacctg gagaacttcc cctacagcca gcagccgctc 1620
115 agcaccgggg ccttcccgcg agggatcact gaccacagcc acttcatgcc cctgctcaat 1680
116 cctcccccaa cggatgccac cagctctgtg gacacccagg ctggcaactg caagccctt 1740
117 cagaaggaca agctccctga gaacctgtg tgggatctca gctgcagag cctcacggcg 1800
118 ctgaccttac aggtggagaa catctccaac accgtccagc agctgctgct ctccaaggct 1860
119 gctgtgccgc agaagaaagg tgtcaagaac ctctgttcca ggacccaga gcagcataaa 1920
120 agccagcact gcagcccgga agggagcggc tactcagccg agcccgagg caccggctg 1980
121 tcagagccgc cagcagcagc gccacaglcc acgcctgcgg agccgcagga ggccgactac 2040

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/508,821A

DATE: 02/23/2001
TIME: 16:03:53

Input Set : A:\Pto.amc
Output Set: N:\CRF3\02232001\I508821A.raw

```

122 ctgagcggct cccaggaccc actggagcgc agcttctctt actgcaacca ggccecgigc 2100
123 agccctgcca gggtaacacg caactcgagc gccaaagccc aglccgtatc cactgtttct 2160
124 gtgacctctc ctgacgacat gtccaccaaa tctgacgact ccttccagag cctacacggc 2220
125 agtctgccgc tccagagctt ctccaaagtc gtggcgggtg agcgggactg tcccgcgctg 2280
126 ctgctcagcg ccttggcaca gtaggacctg gctcgcgaga tccctgggct gcaggaaagc 2340
127 atcgggtgaga agcccgacaa agcltgggcl gaagcaccga gcttggctca ggacagcagc 2400
128 aagccaccct tctcgttggg aaaccacagc gcttgcctgg actctgttgc caaagatgcy 2460
129 tggcccccggc ctggggagcc ggaagccctg cccgactcct tgcagctgga caaggcgccg 2520
130 aatgccaaag acttccagcc agggctgttt gaagaccctt ccttggcctt cgtacgcct 2580
131 gaccccaaaa agacaactgg tctctctctc ttgtgtacca agccaccctt tggggttcct 2640
132 gctccagacc ccaactacgc agctttttgac tgttcccccg acacaaccgc tgcagctca 2700
133 cgggacagcc ccaaccctt tgcctggcca gaggaaaaac tgggggatgc ttgtccagg 2760
134 tggggattgc accctggcga gcttaccagc ggcctggagc aggggtggaa ggcctcagat 2820
135 ggcctcagca aaggggacac ccatgaggtt tgggcttccc tgggcttcca ggaggagac 2880
136 cccctggggg aaaaaglggc clcgttgcgc ggggacttca agcaggagga gglgggtggg 2940
137 gtgaaggagg aggcaggtgg gctgttcagc tgcccagagg tggccaagge tgaccggtgg 3000
138 ctggagggaca ccgggcactg ctgttccacc ggcgacttcc gggacctccc actgctgcca 3060
139 ccaccacaga ggaaggagga cctggaagct gaggaggagt actcctcctt atgtgagctc 3120
140 ctgggcagcc ccgagcagag gcttggcctg caggaccgcg tgtcacccaa ggcctcactc 3180
141 atctgaccca aggaggaggt ggaaggaggt ctggactcca aggcgggctg ggcctctccg 3240
142 tgccacctct caggggagtc cgtcactctg ctgggcctta cagtgggcac cagtcctaac 3300
143 gtcacagact ggttgcagtc ctctcttcca cactgaagc cagggtgaaga ggggctgat 3360
144 ggggagcgag ctccagggga ttccaccacc tgggacgctt ctctggccca gaagcccaac 3420
145 aagcctgtcg tgcccgagge gccatcgcca aagaaagagc ctgtgcacgc gggcaaaagc 3480
146 ttacggagcc gtccgggtgca cggggggctg cccgagggcg aggaactccc atgcagggca 3540
147 ccagtgctgc ccaagacctt cttgtctcct gaatcctgca caggggcccc ccaggggacg 3600
148 atgggaaggcg cgtccagccc aggcgggggg gctcgggaag ggtcctccag gatgttlacl 3660
149 cgttctctca cggccttgag tgagcccgcc acgcccggac ccccaggcct gaccaccacc 3720
150 cctgcacccc cagacaaact gggggggcag cagcgagccg ccttcaagtc gggcaagcgg 3780
151 gtggggaagc cctcacccaa ggcctcctcc agcccagca acccggcgcg cctgcctgtg 3840
152 gctcgcgaca gcagcccgat gggctccagc acaagggaga cagactcacc cagcacgcct 3900
153 ggcaggagcc agcgtcccat gatccttcgg tccagcacca aaaccagga gatcttccac 3960
154 tccaagcggc ggagggcctc tgagggcgcg ctccccaact gccgtgccac caagaagctc 4020
155 ctcgacacaa gccacttgcc cgcacatctc aaggtctcca gcagccccc gaaggagggc 4080
156 aggggtgagc agcgggcaag ggtccccaaa cctggtgcag gcagcaagct ctctgacgg 4140
157 cccctccatg cgtcctaaag gaagtgcgct ttcattggcg cggctccccc caagaagcgg 4200
158 aacctggtct tgccgcagcg cagcagcagc agcagcaacg ccagtgcaat gggggagatg 4260
159 ggaaggagga gaggcctgag ggttccccca cctcttccaa gaggatgtct tctcccaaga 4320
160 aagccaaagc caccaaaggg aatggcgagc ctgcacaaaa gctccacccc ccggagagcc 4380
161 ccattcctgc ctcaagctcg cctctcgcca gcttccagg gggccatgaa gaccaagggt 4440
162 ctgccacccc ggaaggcgcg gggcctgaag ctggaagcca tctgtcagaa gatcacctcg 4500
163 ccagcctca agaagttcgc atgtaaagcg ccaggggcct ctctgttaa tctctgagc 4560
164 ccattccttt ccgacaaaga ccgtgggctc aaggggtgct ggggcagccc agtgggggtg 4620
165 gaagaaggcc ttgtaaatgt gggcaccggg cagaagctcc caacttctgg ggtgatccg 4680
166 tlatgcagaa atccaaacaa cagatcctta aaaggcaaac tcatgaacag taagaaactg 4740
167 tcttctactg actgtttcaa aaccgaggcc ttcacatccc cggaggccct gcagcctggg 4800
168 gggactgccc tggcgctcaa gaaggagagc cggaaaggcc gggcaggggc ccattggactc 4860
169 tccaaaggcc cgttgagaa ggggccttat ctggcccgct gactcccgca 4920
170 gacaggggca gtggcacaac aggggcagat gaggacaact ctgggtggag aggcagaag 4980

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/508,821A

DATE: 02/23/2001
 TIME: 16:03:53

Input Set : A:\Pto.amc
 Output Set: N:\CRF3\02232001\I508821A.raw

```

171 ccaaagatgg aggaagctgg cctgacctc cagcccccgg agggcaggcc ctgcccagccc 5040
172 cagacaaagg cacagaaaca gccaggccac accaactaca gcagctatc caagcgggaag 5100
173 cgcctcactc ggggccgggc caagaaacac acctcttcac cctqtaaggg gcgtgccaaag 5160
174 cgaagacgac aacagcaggt gctgcccctg gatcccgag agcctgaaal ccgcctcaag 5220
175 tacatttctt cttgcaagcg gctgagggtc gacagccgga ccccgccctt ctccaccttc 5280
176 gtgcgggtgg agaagcgaga cgcgttcacc accatalgca ctgltgtcaa ctccctgga 5340
177 gatgcgcccc agcccccacag gaagccttcc tcctctgctt cctcttcttc atcctcgtcc 5400
178 tcgtttctct tggatgcagc cggggccctc ctggccacac tccttgagg ctccatctctg 5460
179 cagccgcggc cctccttgcc cctctctccc acgatgcact tggggcctgt ggtttccaaag 5520
180 gccctgagta cctcttgctt tgtttgctgc ctctgcaaaa acccgcccaa ctccaaggac 5580
181 cttggggacc tctgtgggcc ctactacct gaacactgcc tccccaaaaa gaagccaaaa 5640
182 ctcaaggaga aggtgcggcc agaaggcacc tgtgaggagg cctcgtctgc gcltgagaga 5700
183 aactcacaag gtcccgagtg tgcagctgcc gccactgcc ggaagcctcc caggtgacgg 5760
184 ccagctgac ccggccaagc agggccact gcgcaccagt gcccggggcc tgtcccgag 5820
185 gctgcagagc tgcctactgt gtgctggccg ggaggatgg ggcgaggagg cagcccccagc 5880
186 cgaacaaggt cgaacaatg agtgacagca ggaggctccg gcagagcccg gcggggaggc 5940
187 ccaggagcac tgggtgcatg aggcctgtgc cgtgtggacc ggcgcgctct acctggtggc 6000
188 cgggaagctc ttggyctgc ag                                     6022

191 <210> SEQ ID NO: 6
192 <211> LENGTH: 22
193 <212> TYPE: DNA
194 <213> ORGANISM: Artificial Sequence
196 <220> FEATURE:
197 <223> OTHER INFORMATION: Description of Artificial Sequence:
198     oligonucleotide
200 <400> SEQUENCE: 6
201 tggccttget gccgtagtg ct

```

22

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/508,821A

DATE: 02/23/2001
TIME: 16:03:54

Input Set : A:\Pto.amc
Output Set: N:\CRF3\02232001\I508821A.raw

L:48 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2